Our Ref.: 3660-27
P1874-135 US JA/vo

U.S. PATENT APPLICATION

Inventor(s): Alberto JIMENEZ FELSTRÖM
Jim RASMUSSON

Invention: Weighted Spectral Distance Calculator

NIXON & VANDERHYE P.C.
ATTORNEYS AT LAW
1100 NORTH GLEBE ROAD
8TH FLOOR
ARLINGTON, VIRGINIA 22201-4714
(703) 816-4000
Facsimile (703) 816-4100

SPECIFICATION

APPLICANT: TELEFONAKTIEBOLAGET L M ERICSSON
TITLE: WEIGHTED SPECTRAL DISTANCE CALCULATOR

Field of the Invention

The present invention relates to a spectral distance calculator, and more particularly to a spectral distance calculator, comprising means for performing spectral distance calculations for comparison of an input spectrum, in the presence of noise, and a reference spectrum.

Description of the Prior Art

Speech recognition systems can be used to enter data and information in order to control different kinds of electronic apparatuses. Despite some limitations, speech recognition has a number of applications, for example mobile phones can be provided with automatic speech recognition functionalities, in particular so called automatic voice answering (AVA) functions.

SUB B2

An example of an AVA function is the possibility to accept or reject an incoming call to a mobile phone by using the voice instead of a manual activation through for example a key stroke on the key pad of the phone. Such a function is applicable for a user of a mobile phone when he is for example driving a vehicle. When the user of the phone drives his vehicle and the mobile phone indicates an incoming call by a ring signal the user can give speech commands to control the phone.

A problem associated with AVA functions is that the ring signal emitted by the phone interferes strongly with the given AVA command.

Some prior art phones are provided with a simple kind of AVA functionality based on energy detectors. The phone is responsive and detects an AVA command when the speech has a higher energy level than a pre-defined threshold. As

a result, only one answering function can be provided, usually "reject the call" is chosen.

A state of the art mobile phone provided by the applicant, Ericsson T18, is provided with an automatic voice dialling function.

AVA functions based on energy detectors are restrained to accept only one command as mentioned above. It is not convenient to provide several commands when the AVA functions are based on energy detectors, because it is very likely that the AVA functions of the phone will response to sounds like the ring signal surrounding the phone.

SUB B3

On the other hand, AVA functions based on speech recognition are sensitive to interference from other sounds having similar spectral characteristics as speech. One reason for this is that the dissimilarity measure used in the speech recognizer is mostly based on the difference between short time spectral characteristics of the perceived sound or speech and of the pre-trained speech references or templates.

Another solution is based on low-pass filtering of the microphone signals which increases the recognition rate of the AVA commands. However, a disadvantage of this solution is that all speech information having frequencies above the filter cut-off frequency cannot be used by a speech recognizer even though the ring signal does not cover all frequencise above the cut-off.

In still another approach to solve this problem, the mobile phone can be provided with an adaptive filter between the microphone and the speech recognizer in order to filter out different ring signals.

The adaptive filter can be interpreted as an adaptive notch filter, wherein the location of the notches are updated continuously in a way that only disturbed frequencies are attenuated. As a result higher recognition

SUB B4

SUB B4

rates are achieved by using this method. However, such adaptive algorithm needs a lot of calculations. Further, they do not adapt instantaneously and a trade off between stability and the convergence time for the adaptation have to be performed.

GB-A-2 137 791 discloses a spectral distance processor for comparing spectra taken from speech in the presence of background noise which has to be estimated. In order to prepare an input spectrum and a template spectrum for comparison, the processor includes means for masking the input spectrum with respect to an input noise spectrum estimate, means for masking the template spectrum with spectrum to a template noise spectrum estimate, and means for marking samples of each masked spectrum dependant upon whether the sample is due to noise or speech.

SUB B5

During the masking operations noise marks are associated with the masked input spectrum and template spectrum, respectively, whether the value arose from noise or speech and taken into accont during spectral distance calculations on the spectra.

Where the greater of the masked spectral samples is marked to be due to noise, a default noise distance is assigned in place of the distance between the two masked spectra.

SUB B6

Hence, since the spectral distance processor according to GB-A-2 137 791 is intended to operate in fluctutating or high noise level conditions that's the reason for the complex design.

SUB B7

However, speech recognition in a mobile phone where the user can give speech commands to control the phone as described above, a complex spectral distance processor as disclosed by GB-A-2 137 791 is not necessary, because the present noise dose not fluctuate and has no such high level.

Summary of the Invention

Therefore, it is an object of the present invention to provide an improved spectral distance calculator usable in any speech recognition using spectral difference as a dissimilarity measure, particularly suitable in low noise level conditions.

SubB8

In accordance with one aspect of the present invention, the spectral distance calculator comprises spectral distance calculation means for performing spectral distance calculations for comparison of an input spectrum, from an input signal in the presence of a noise signal, and a reference spectrum, memory means for pre-storing a noise spectrum from the noise signal, and means for masking the spectral distance between the input spectrum and the reference spectrum with respect to the pre-stored noise spectrum.

In accordance with another aspect of the present invention, the noise has a lower level than the input spectrum.

Another object of the invention is to provide a speech recognition system for comparison of an input spectrum and a reference spectrum including a spectral distance calculator as mentioned above, wherein the recognition system comprises selecting means for selecting a reference spectrum minimizing a complete spectral distance between the input spectra and the reference spectra.

Still another object of the invention is to provide a mobile phone including the speech recognition system as described.

An advantage of the present invention is that automatic voice answering functions (AVA) of a mobile phone having a speech recognition system, provided with a spectral distance calculator according to the invention, is

reliable in responding to different AVA commands in presence of ring signals surrounding the phone.

Brief Description of the Drawings

In order to explain the invention in more detail and the advantages and features of the invention a preferred embodiment will be described in detail below, reference being made to the accompanying drawings, in which

FIG 1 shows an example of an input spectra due to a known noise, a reference spectra, and the known noise spectra; and

FIG 2 illustrates the noise compensation according to the invention.

Detailed Description of the Invention

One embodiment of a spectral distance calculator according to the invention comprises spectral distance calculation means for performing spectral distance calculations in order to compare an input spectrum due to noise and a reference spectrum. In order to deal with the interfering noise, the distance calculator further comprises masking means in order to mask the spectral distance between the input spectrum and the reference spectrum with respect to a known or pre-defined noise, stored in a memory means.

The distance calculator in the embodiment is based on city distances and discrete spectral representation of a speech. However, this solution can be generalized to other spectral representation of the speech within the scope of the invention.

Further, a spectral distance calculator according to the invention can be used in any speech recognition system, using spectral difference as dissimilarity or distance measure, for example in a mobile phone controlled by speech commands.

A user of a speech recognition system speaks into a microphone, wherein each sound is broken down into its various frequencies. The received sounds in each frequency are digitized so they can be manipulated by the speech recognition system. The microphone signal is denoted by s(n) and its corresponding spectral representation is denoted by $S_n(f)$, where n is the time for each sample and f is the current frequency.

Sub B10

The digitized version of the sound is matched against a set of templates or reference signals pre-stored in a system storage. A template or reference signal is denoted by r(n) and a corresponding spectral representation of the template signal is denoted by $R_n(f)$. The known noise signal in the input is denoted by x(n) and the corresponding spectral representation is denoted by $X_n(f)$.

The measure of the dissimilarity or distance used in a speech recognizer is for example given by the expression:

$$D_n = \sum_i \left| R_n(f_i) - S_n(f_i) \right|$$

Sub B11

Thus, the input signal spectrum $S_n(f)$ is matched against similarly formed reference signals $R_n(f)$ among the stored reference signals in the electronic storage. This match procedure is performed by selecting the reference signal which minimizes the complete spectral distance, i.e is minimizing the following expression:

$$\sum_n D_n$$

However, this selection procedure does not take into consideration any information about interfering noise signals.

In a mobile phone providing speech recognition functions or in particular so called automatic voice answering (AVA) functions the ring signal emitted by the phone interferes strongly with the given AVA command.

The ring signal is a known "noise" signal and, consequently, the spectrum representing the ring signal can be pre-stored in the memory means associated with the spectral distance calculator.

The ring signal is for example a buzzer or a personal ring signal, such as a simple melody, selected or programmed by the user. However, when the ring signal is selected or programmed it is "known" by the phone and a spectrum representing the current ring signal can be stored in the memory means for pre-stored noise spectra. In an alternative embodiment a plurality of spectra from different ring signals can be pre-stored and the current ring signal selected is marked by a bit set in the memory. Then, the spectral distance calculator can identify and select the current spectrum to be used in the masking procedure according to the invention.

According to fig 1, the input signal for a comparison is exposed to a known noise in the spectrum between the two frequencies f_a and f_b. The corresponding reference signal $R_n(f)$ for comparison with the input signal is not considered to be due to any noise. Hence, in order to get a thorough comparison between the input signal and the reference signal or their spectra, the input signal has to be masked in any way to compensate for the known noise. According to the invention, the spectral distance calculation or measure of the dissimilarity is modified by a weight A_i according to the following expression:

$$D_n = \sum_i A_i \left| R_n(f_i) - S_n(f_i) \right| .$$

In this expression A_i is equal to zero if the frequency f_i of the input signal is due to any known noise and A_i is unity if no noise is present at the current frequency f_i.

FIG 2 illustrates the noise compensation according to the invention, wherein the spectral distance between the

input spectrum $S_n(f_i)$ and the reference spectrum $R_n(f_i)$ is assigned a zero value in the spectrum between the two frequencies f_a and f_b.

SUBB13

In one embodiment of the spectral distance calculator according to the invention it is included in a speech recognition system for comparison of an input spectrum and a reference spectrum, comprising selecting means for selecting a reference spectrum minimizing a complete spectral distance between the input spectra and the reference spectra.

SUBB14

Further, the speech recognition system is included in a mobile phone providing AVA functions, such as "accept the call" if a user of the phone would like to answer the call, or "reject the call" if he doesn't want to answer the call, or "forward" if the incoming call should be connected to a voice mail or another phone number.

SUBB15

Although the invention has been described by way of a specific embodiment thereof, it should be apparent that the present invention provides a weighted spectral distance calculator that fully satisfies the aims and advantages set forth above, and alternatives, modifications and variations are apparent to those skilled in the art.

SUBB16

For example, in another embodiment of the invention the calculator is provided with an adaptive notch filter which not only filters the input signal but also the reference signal. This solution benefits from the effect that a more reliable selection of the reference signal is obtained because the calculation will be more accurate if a filtered input signal is compared to a filtered reference signal. Further, this solution does not require any adaptive algorithms and there is no additional computational loading, it works instantaneously and it lacks stability problems. However, the automatic voice answering means requires continuously knowledge of the disturbed frequencies.

SUBB17

In alternative embodiments of the second embodiment, more sophisticated weights are provided by using real valued A_i, allowing different levels of suppression depending on how much the specific frequencies f_i are disturbed.

6. A mobile telephone including a speech recognition system according to claim 4 ~~or 5~~, characterized by call answering means operatively connected to said speech recognition system, wherein said answering means is responsive to speech answering commands.

7. A mobile telephone according to claim 6, characterized in that said answering means is responsive to an accept call command for accepting a call.

8. A mobile telephone according to claim 6 ~~or 7,~~ characterized in that said answering means is responsive to a reject call command for rejecting a call.

9. A mobile telephone according to claim ~~any of the claims~~ 6~~-8~~, characterized in that said answering means is responsive to a forward call command for forwarding a call.